



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/071,838

DATE: 05/16/2002 16
TIME: 13:06:54

Input Set : A:\-76-1.app

Output Set: N:\CRF3\05162002\J071838.raw

3 <110> APPLICANT: Li, Jing
 4 Powers, Scott
 5 Xiang, Phil
 6 Peng, Yue
 7 Tularik Inc.
 9 <120> TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
 11 <130> FILE REFERENCE: 018781-007610US
 13 <140> CURRENT APPLICATION NUMBER: US 10/071,838
 C--> 14 <141> CURRENT FILING DATE: 2002-05-07
 16 <150> PRIOR APPLICATION NUMBER: US 60/267,615
 17 <151> PRIOR FILING DATE: 2001-02-08
 19 <160> NUMBER OF SEQ ID NOS: 18
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 1964
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Homo sapiens
 28 <220> FEATURE:
 29 <223> OTHER INFORMATION: human PRC17
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (1)..(1650)
 35 <400> SEQUENCE: 1
 36 atggacgtgg tagaggtcgc gggcagttgg tgggcacaag agcgagagga catcattatg 60
 37 aaatacga aaaggacaccg agctgggctg ccagaggaca aggggcctaa gccttttcga 120
 38 agctacaaca acaacgtcga tcatattgggg attgtacatg agacggagct gcctcctctg 180
 39 actgcgcggg aggcgaagca aattcggcgg gagatcagcc gaaagagcaa gtgggtggat 240
 40 atgctgggag actgggagaa atacaaaagc agcagaaagc tcatagatcg agcgtacaag 300
 41 ggaatgcccc tgaacatccg gggcccgatg tggtcagtcc tctgaacat tgaggaaatg 360
 42 aagttgaaaa accccggaag ataccagatc atgaaggaga agggcaagag gtcattctgag 420
 43 cacatccagc gcatcgaccg ggacgtaagc gggacattaa ggaagcatat attcttcagg 480
 44 gatcgatacg gaaccaagca gcgggaacta ctccacatcc tctggcata tgaggagtat 540
 45 aaccgcggagg tgggctactg cagggacctg agccacatcg ccgccttggt cctcctctat 600
 46 ctctctgagg aggatgcatt ctgggcaactg gtgcagctgc tggccagtga gaggcaactc 660
 47 ctgcagggat ttacagccc aaatggcggg accgtccagg ggctccaaga ccaacaggag 720
 48 catgtggtag ccacgtcaca acccaagacc atggggcatc aggacaagaa agatctatgt 780
 49 gggcagtggt ccccgttagg ctgcctcatc cggatattga ttgacgggat ctctctcggg 840
 50 ctcaccctgc gcctgtggga cgtgtatctg gtagaaggcg aacaggcggt gatgccgata 900
 51 acaagaatcg cctttaaggt tcagcagaag cgcctcacga agacgtccag gtgtggcccc 960
 52 tgggcacggt ttgcaaccg gttcgttgat acctgggcca gggatgagga cactgtgctc 1020
 53 aagcatctta gggcctctat gaagaaacta acaagaaagc agggggacct gccaccccca 1080
 54 gccaaacccg agcaagggtc gtccggcatcc aggcctgtgc cggcttcacg tggcgggaag 1140
 55 accctctgca agggggacag gcaggccctc ccaggccac cagcccggtt cccgcggccc 1200

RAW SEQUENCE LISTING

DATE: 05/16/2002

PATENT APPLICATION: US/10/071,838

TIME: 13:06:54

Input Set : A:\-76-1.app

Output Set: N:\CRF3\05162002\J071838.raw

```

56 atttgggtcag cttccccgcc acggggcacct cgtttcttcca caccctgtcc tgggtggggct 1260
57 gtccggggaag acacctaccc tgtgggcact caggggtgtgc ccagcccggc cctgggtcag 1320
58 ggaggacctc agggttcctg gagattcctg cagtgggaact ccattgcccc cctcccaacg 1380
59 gacctggacg tagaggggccc ttgggttccgc cattatgatt tcagacagag ctgctgggtc 1440
60 cgtgccatat cccaggagga ccagctggcc ccctgctggc aggtgaaca ccctgcggag 1500
61 cgggtgagat cggcttctgc tgcacccagc actgattccg accagggcac ccccttcaga 1560
62 gctagggacg aacagccgtg tgctcccacc tcagggcctt gcctctgcgg cctccacttg 1620
63 gaaagtcttc agttccctcc aggtctctag aagcatctgg gccagggctc atggctggat 1680
64 aatttcccta ggcttaacaa cccaagcaag cttcgcgtcc tcgttttatt ttgggttaaa 1740
65 cttatgaaaa tgtattaaga aagagtgcag ctcgagagag attcagagat ggaacacacc 1800
66 agaccccaga tcacaaagcc aaccatgccc agcccctccc agcaccocca gcccacgac 1860
67 catcgttctg aattctgacg acaccgtgag cctgcctttg tactttaaac tcattggaag 1920
68 ataactacct tcacgttttg aaataaatgt ttcctgttga aatg 1964

```

71 <210> SEQ ID NO: 2

72 <211> LENGTH: 549

73 <212> TYPE: PRT

74 <213> ORGANISM: Homo sapiens

76 <220> FEATURE:

77 <223> OTHER INFORMATION: human PRC17

79 <400> SEQUENCE: 2

```

80 Met Asp Val Val Glu Val Ala Gly Ser Trp Trp Ala Gln Glu Arg Glu
81   1           5           10           15
83 Asp Ile Ile Met Lys Tyr Glu Lys Gly His Arg Ala Gly Leu Pro Glu
84           20           25           30
86 Asp Lys Gly Pro Lys Pro Phe Arg Ser Tyr Asn Asn Asn Val Asp His
87           35           40           45
89 Leu Gly Ile Val His Glu Thr Glu Leu Pro Pro Leu Thr Ala Arg Glu
90           50           55           60
92 Ala Lys Gln Ile Arg Arg Glu Ile Ser Arg Lys Ser Lys Trp Val Asp
93   65           70           75           80
95 Met Leu Gly Asp Trp Glu Lys Tyr Lys Ser Ser Arg Lys Leu Ile Asp
96           85           90           95
98 Arg Ala Tyr Lys Gly Met Pro Met Asn Ile Arg Gly Pro Met Trp Ser
99           100          105          110
101 Val Leu Leu Asn Ile Glu Glu Met Lys Leu Lys Asn Pro Gly Arg Tyr
102           115          120          125
104 Gln Ile Met Lys Glu Lys Gly Lys Arg Ser Ser Glu His Ile Gln Arg
105           130          135          140
107 Ile Asp Arg Asp Val Ser Gly Thr Leu Arg Lys His Ile Phe Phe Arg
108 145           150           155           160
110 Asp Arg Tyr Gly Thr Lys Gln Arg Glu Leu Leu His Ile Leu Leu Ala
111           165           170           175
113 Tyr Glu Glu Tyr Asn Pro Glu Val Gly Tyr Cys Arg Asp Leu Ser His
114           180           185           190
116 Ile Ala Ala Leu Phe Leu Leu Tyr Leu Pro Glu Glu Asp Ala Phe Trp
117           195           200           205
119 Ala Leu Val Gln Leu Leu Ala Ser Glu Arg His Ser Leu Gln Gly Phe
120           210           215           220
122 His Ser Pro Asn Gly Gly Thr Val Gln Gly Leu Gln Asp Gln Gln Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/071,838

DATE: 05/16/2002

TIME: 13:06:54

Input Set : A:\-76-1.app

Output Set: N:\CRF3\05162002\J071838.raw

```

123 225          230          235          240
125 His Val Val Ala Thr Ser Gln Pro Lys Thr Met Gly His Gln Asp Lys
126          245          250          255
128 Lys Asp Leu Cys Gly Gln Cys Ser Pro Leu Gly Cys Leu Ile Arg Ile
129          260          265          270
131 Leu Ile Asp Gly Ile Ser Leu Gly Leu Thr Leu Arg Leu Trp Asp Val
132          275          280          285
134 Tyr Leu Val Glu Gly Glu Gln Ala Leu Met Pro Ile Thr Arg Ile Ala
135          290          295          300
137 Phe Lys Val Gln Gln Lys Arg Leu Thr Lys Thr Ser Arg Cys Gly Pro
138 305          310          315          320
140 Trp Ala Arg Phe Cys Asn Arg Phe Val Asp Thr Trp Ala Arg Asp Glu
141          325          330          335
143 Asp Thr Val Leu Lys His Leu Arg Ala Ser Met Lys Lys Leu Thr Arg
144          340          345          350
146 Lys Gln Gly Asp Leu Pro Pro Pro Ala Lys Pro Glu Gln Gly Ser Ser
147          355          360          365
149 Ala Ser Arg Pro Val Pro Ala Ser Arg Gly Gly Lys Thr Leu Cys Lys
150          370          375          380
152 Gly Asp Arg Gln Ala Pro Pro Gly Pro Pro Ala Arg Phe Pro Arg Pro
153 385          390          395          400
155 Ile Trp Ser Ala Ser Pro Pro Arg Ala Pro Arg Ser Ser Thr Pro Cys
156          405          410          415
158 Pro Gly Gly Ala Val Arg Glu Asp Thr Tyr Pro Val Gly Thr Gln Gly
159          420          425          430
161 Val Pro Ser Pro Ala Leu Ala Gln Gly Gly Pro Gln Gly Ser Trp Arg
162          435          440          445
164 Phe Leu Gln Trp Asn Ser Met Pro Arg Leu Pro Thr Asp Leu Asp Val
165          450          455          460
167 Glu Gly Pro Trp Phe Arg His Tyr Asp Phe Arg Gln Ser Cys Trp Val
168 465          470          475          480
170 Arg Ala Ile Ser Gln Glu Asp Gln Leu Ala Pro Cys Trp Gln Ala Glu
171          485          490          495
173 His Pro Ala Glu Arg Val Arg Ser Ala Phe Ala Ala Pro Ser Thr Asp
174          500          505          510
176 Ser Asp Gln Gly Thr Pro Phe Arg Ala Arg Asp Glu Gln Pro Cys Ala
177          515          520          525
179 Pro Thr Ser Gly Pro Cys Leu Cys Gly Leu His Leu Glu Ser Ser Gln
180          530          535          540
182 Phe Pro Pro Gly Phe
183 545
186 <210> SEQ ID NO: 3
187 <211> LENGTH: 2146
188 <212> TYPE: DNA
189 <213> ORGANISM: Homo sapiens
191 <220> FEATURE:
192 <223> OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
194 <220> FEATURE:
195 <221> NAME/KEY: CDS

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/071,838

DATE: 05/16/2002

TIME: 13:06:54

Input Set : A:\-76-1.app

Output Set: N:\CRF3\05162002\J071838.raw

```

196 <222> LOCATION: (1)..(1832)
198 <400> SEQUENCE: 3
199 atggacgtgg tagaggtcgc gggcagttgg tgggcacaag agcgagagga catcattatg 60
200 aaatacgaag agggacaccg agctgggctg ccagaggaca aggggcctaa gcccttttca 120
201 agctacaaca acaacgtcga tcatttgggg attgtacagt cctgccgctc ctgggagtcg 180
202 gccccacagg aaggcccttg tctcccttc cctgtgcctt ctcctgggct gagccctgag 240
203 ctggaaaagg acagagccag tcccttctgg gggctggcac ccaggctggg gccgctccag 300
204 gccccgtgca gttcctcagc tctgcctggg ttgccttaca gtgagacgga gctgcctcct 360
205 ctgactgcgc gggaggcgaa gcaaattcgg cgggagatca gccgaaagag caagtgggtg 420
206 atatgctggg agactgggag aaatacaaaa gcagcagaaa gctcatagat cgagcgtaca 480
207 aggggaatgcc catgaacatc cggggcccgga tgtggtcagt cctcctgaac attgaggaaa 540
208 tgaagttaa aaaccccgga agataccaga tcatgaagga gaagggcaag aggtcatctg 600
209 agcacatcca gcgcctcagc cgggacgtaa gcgggacatt aaggaagcat atattcttca 660
210 gggatcgata cggaaccaag cagcgggaac tactccacat cctcctggca tatgaggagt 720
211 ataaccggga ggtgggctac tgcagggacc tgagccacat cgccgccttg ttctcctct 780
212 atcttctga ggaggatgca ttctgggcac tgggtgcagt gctggccagt gagaggcact 840
213 cctgcaggg atttcacagc ccaaattggc ggaccgtcca ggggctccaa gaccaacagg 900
214 agcatgtggt agccacgtca caaccaaga ccatggggca tcaggacaag aaagatctat 960
215 gtgggcagtg ttccccgtta ggctgcctca tccggatatt gattgacggg atctctctcg 1020
216 ggctcaccct gcgcctgtgg gacgtgtatc tggtagaagg cgaacaggcg ttgatgccga 1080
217 taacaagaat cgcctttaag gttcagcaga agcgcctcac gaagacgtcc aggtgtggcc 1140
218 cgtgggcacg tttttgcaac cgggttcgtt atacctgggc cagggatgag gacactgtgc 1200
219 tcaagcatct tagggcctct atgaagaaac taacaagaaa gcagggggac ctgccacccc 1260
220 cagccaaaacc cgagcaaggg tgcctggcat ccaggcctgt gccggcttca cgtggcgagg 1320
221 agaccctctg caagggggac aggcaggccc ctccaggccc accagcccgg ttcccgcggc 1380
222 ccatttggtc agcttccccg ccacgggcac ctcggttctc cacacctgt cctgggtggg 1440
223 ctgtccggga agacacctac cctgtgggca ctccagggtg gccagcccgg gccctggctc 1500
224 agggaggacc tcagggttcc tggagattcc tgcagtggaa ctccatgcc cgctcccaa 1560
225 cggacctgga cgtagagggc ccttggttcc gccattatga ttccagacag agctgctggg 1620
226 tccgtgccat atcccaggag gaccagctgg cccctgctg gcaggctgaa caccctgcgg 1680
227 agcgggtgag atcggtttc gctgcaccca gcactgattc cgaccagggc acccccttca 1740
228 gagctaggga cgaacagccg tgtgctccca cctcagggcc ttgcctctgc ggccctcact 1800
229 tggaaaagttc tcagttccct ccaggcttct agaagcatct gggccagggc tcatggctgg 1860
230 ataatttccc taggttaaac aacccaagca agcttcgcgt cctcgtttta tttttgggta 1920
231 aacttatgaa aatgtattaa gaaagagtgc agctcgagag agattcagag atggaacaca 1980
232 ccagacccca gatcacaag ccaacctatg ccagcccctc ccagcaccac cagcccccag 2040
233 accatcggtc tgaattctga cgacaccgtg agcctgcctt tgtactttaa actcatggaa 2100
234 ggataactac cttcacgttt tgaataaat gtttctgtt gaaatg 2146
237 <210> SEQ ID NO: 4
238 <211> LENGTH: 610
239 <212> TYPE: PRT
240 <213> ORGANISM: Homo sapiens
242 <220> FEATURE:
243 <223> OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
245 <400> SEQUENCE: 4
246 Met Asp Val Val Glu Val Ala Gly Ser Trp Trp Ala Gln Glu Arg Glu
247 1 5 10 15
249 Asp Ile Ile Met Lys Tyr Glu Lys Gly His Arg Ala Gly Leu Pro Glu
250 20 25 30

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/071,838

DATE: 05/16/2002

TIME: 13:06:54

Input Set : A:\-76-1.app

Output Set: N:\CRF3\05162002\J071838.raw

```

252 Asp Lys Gly Pro Lys Pro Phe Arg Ser Tyr Asn Asn Asn Val Asp His
253          35          40          45
255 Leu Gly Ile Val Gln Ser Cys Arg Ser Trp Glu Ser Ala Pro Gln Glu
256          50          55          60
258 Gly Pro Cys Pro Pro Phe Pro Val Pro Ser Pro Gly Leu Ser Pro Glu
259 65          70          75          80
261 Leu Glu Arg Asp Arg Ala Ser Pro Phe Trp Gly Ser Ala Pro Arg Leu
262          85          90          95
264 Gly Pro Leu Gln Ala Pro Cys Ser Ser Ser Ala Leu Pro Gly Leu Pro
265          100          105          110
267 Tyr Ser Glu Thr Glu Leu Pro Pro Leu Thr Ala Arg Glu Ala Lys Gln
268          115          120          125
270 Ile Arg Arg Glu Ile Ser Arg Lys Ser Lys Trp Val Asp Met Leu Gly
271          130          135          140
273 Asp Trp Glu Lys Tyr Lys Ser Ser Arg Lys Leu Ile Asp Arg Ala Tyr
274 145          150          155          160
276 Lys Gly Met Pro Met Asn Ile Arg Gly Pro Met Trp Ser Val Leu Leu
277          165          170          175
279 Asn Ile Glu Glu Met Lys Leu Lys Asn Pro Gly Arg Tyr Gln Ile Met
280          180          185          190
282 Lys Glu Lys Gly Lys Arg Ser Ser Glu His Ile Gln Arg Ile Asp Arg
283          195          200          205
285 Asp Val Ser Gly Thr Leu Arg Lys His Ile Phe Phe Arg Asp Arg Tyr
286          210          215          220
288 Gly Thr Lys Gln Arg Glu Leu Leu His Ile Leu Leu Ala Tyr Glu Glu
289 225          230          235          240
291 Tyr Asn Pro Glu Val Gly Tyr Cys Arg Asp Leu Ser His Ile Ala Ala
292          245          250          255
294 Leu Phe Leu Leu Tyr Leu Pro Glu Glu Asp Ala Phe Trp Ala Leu Val
295          260          265          270
297 Gln Leu Leu Ala Ser Glu Arg His Ser Leu Gln Gly Phe His Ser Pro
298          275          280          285
300 Asn Gly Gly Thr Val Gln Gly Leu Gln Asp Gln Gln Glu His Val Val
301          290          295          300
303 Ala Thr Ser Gln Pro Lys Thr Met Gly His Gln Asp Lys Lys Asp Leu
304 305          310          315          320
306 Cys Gly Gln Cys Ser Pro Leu Gly Cys Leu Ile Arg Ile Leu Ile Asp
307          325          330          335
309 Gly Ile Ser Leu Gly Leu Thr Leu Arg Leu Trp Asp Val Tyr Leu Val
310          340          345          350
312 Glu Gly Glu Gln Ala Leu Met Pro Ile Thr Arg Ile Ala Phe Lys Val
313          355          360          365
315 Gln Gln Lys Arg Leu Thr Lys Thr Ser Arg Cys Gly Pro Trp Ala Arg
316          370          375          380
318 Phe Cys Asn Arg Phe Val Asp Thr Trp Ala Arg Asp Glu Asp Thr Val
319 385          390          395          400
321 Leu Lys His Leu Arg Ala Ser Met Lys Lys Leu Thr Arg Lys Gln Gly
322          405          410          415
324 Asp Leu Pro Pro Pro Ala Lys Pro Glu Gln Gly Ser Ser Ala Ser Arg

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/071,838

DATE: 05/16/2002
TIME: 13:06:55

Input Set : A:\-76-1.app
Output Set: N:\CRF3\05162002\J071838.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; N Pos. 1,20
Seq#:12; N Pos. 1,29

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/071,838

DATE: 05/16/2002

TIME: 13:06:55

Input Set : A:\-76-1.app

Output Set: N:\CRF3\05162002\J071838.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:567 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
L:618 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0